Belief Revision in a Changing World

Robert Koons and Nicholas Asher
Center for Cognitive Science
The University of Texas@Austin
rkoons@emx.cc.utexas.edu
asher@gris.irit.fr

Introduction.

Several authors (Keller and Winslett 1985, Winslett 1988, Katsuno and Mendelzon 1989, Morreau and Rott 1991) have recently argued for a distinction in the way beliefs are updated with new information. They distinguish between information that tells the agent that the world has changed over time and information that fills in or corrects the agent's picture of the world at a particular time. We provide an explicit representation of this distinction by means of a modal logic that combines epistemic and dynamic features. Furthermore, we develop a completely declarative semantics for belief revision. This semantics enables us to deduce the result of revising a given body of beliefs in the light of new information, given simply the semantic content of the prior beliefs and of the new data. No purely procedural assumptions about the agent's epistemic policies or values (no information about priorities of defaults or degrees of entrenchment) are needed, beyond what is explicitly represented in the objects of the agent's beliefs. We accomplish this by distinguishing hard (incorrigible, unrevisable) belief and soft belief; further, the soft attitudes supervene on the hard level. We use a specific theory of nonmonotonic inference to generate soft attitudes from hard ones. This last point is especially important in attempting to deal with belief change, because when an agent acquires new beliefs there is the question: what beliefs about the world persist? We think that only a nonmonotonic logic can adequately deal with this question in a sufficiently rich framework for belief revision like the one we propose.

To motivate this last claim, let us look at what belief change involves in a modal setting. Moore (1980) gives the classic theory of belief revision under the assumption of change in a modal setting. As in modal dynamic logic, Moore analyzed actions, or more generally changes, as state transition relations. If an agent's belief state is represented as a set of possible worlds σ, the effects on the belief state of the agent's having observed the change R is $\mathrm{Rng}(R\restriction\sigma)$. We can then use dynamic logic to reason about belief change. Suppose that the agent's beliefs are characterized in the language of dynamic logic by: $\psi \,\&\, [\alpha]\varphi$, and that the agent observes action α. Moore's theory and the semantics of dynamic logic tell us that φ holds in the updated belief state. But should ψ hold in the updated state? This is the question of information persistence. And neither dynamic logic alone nor Moore's semantic picture answers it.

One paradigm for studying belief dynamics in AI and philosophy currently is the belief *revision* procedure developed by Gärdenfors.[1] However, Gärdenfors's notion of revision is inappropriate for updating beliefs under the assumption of change, as the following very simple example shows:

EXAMPLE 1: THE AUSTIN MERCADO

> Suppose you believe that a basket you are looking at has tomatillos or chiles in it but not both. Now you observe someone putting tomatillos in the basket. What do you conclude?

Suppose now that you revise or update your beliefs in the general way described by Gärdenfors.

DEFINITION OF REVISION:

$$\mathrm{Revision}(\mathrm{Th}(\sigma), \varphi) = \mathrm{Th}(\mathrm{Contract}(\mathrm{Th}(\sigma), \neg\varphi) \cup \{\varphi\})$$

This means that you take away all the information that entails $\neg$(in basket) and you then add to your beliefs the proposition that tomatillos are in the basket. Since beliefs are closed under logical

[1] In AI the Gärdenfors definitions seem to be taken up explicitly in the work of Rao and Foo (1989).

consequence in this approach as well, you conclude that the basket does not have chiles in it. But this seems to be incorrect; this reasoning fails to take account of the way the world has changed.[2] When we try to complicate Gärdenfors's language and add temporal parameters to it so that every propositional formula is understood to hold at a particular time, we no longer make the undesired prediction. Gfrdenfors's theory then predicts only that tomatillos will be in the basket at the time after the tomatillos were added to the basket. But we do so at a considerable cost; we lose what little account of information persistence his definition of belief revision gave us. Suppose that next to the basket you observe a tree prior to the tomatillos being added. We would like to infer that there continues to be a tree next to the basket after the tomatillos have been added, but we cannot do this once temporal parameters have been assigned to formulas. There is no way on this approach to make Tree-next-to-basket(t+1) true after the revision.

This example also illustrates a difficulty with using default logic to represent persistence. If there is a default rule of persistence that applies to all formulas, then in example #1 we will draw the erroneous conclusion that there are no chiles in the basket after the action is performed, applying the default rule of persistence to the biconditional ($\alpha \leftrightarrow \neg\beta$). Alternatively, if we restrict the default rule to literals, we will be unable to account for the fact that disjunctive facts typically do persist, so long as nothing disturbs any of the constituent disjuncts. For example, we will be unable to draw the plausible inference that there would still be either chiles or tomatillos in the basket after performing the action Wait.

Winslett (1988) and Katsuno and Mendelzon (1989) (1991) provide alternative rules of belief revision that do not succumb to the difficulties noted with this particular example. However, Winslett's proposal (which Katsuno and Mendelzon (1991) adopt), on which the update of a cognitive state σ with φ is the state (or intersection of states) which verifies φ and minimizes at each world in the state the set of changes to the truth values of atomic statements (this is what in effect her proposal says which she describes using models), is seriously flawed. It depends on a linguistic distinction (between atomic sentences and others) that we can always circumvent by introducing definitions. For instance let the propositional variable p stand for *there are not both tomatillos and chiles in the basket*. Now if we try to minimize the number of changes in truth value of the atomic statements on updating with the observation that tomatillos are in the basket, we have two states with minimal changes to the truth values of atomic statements, one in which at each world we change the truth value of p and one in which we make q false, where q:= there are chiles in the basket. Further, there are other examples where minimizing the number of changes in truth values to the number of atoms at each world simply goes wrong, as Lifschitz (1987) has argued.

EXAMPLE 2 : THE LIFSCHITZ LAMP

> There are two switches connected to a lamp. When the switches are both down or both up, the light is on; when they are in different positions, the light is off. Furthermore, a switch is up just in case it is not down. Suppose switch 1 is up, switch 2 is down and that you move switch 1 and that moving a switch changes its state.

The atoms of the theory here are simply Up(sw1) and Dn(sw2). According to Winslett's revision procedure, the minimal change in the atoms needed to accommodate the effects of the action is not to make any changes in the truth values of the atoms! Switch 1 stays up, switch 2 stays down, and the rule about the effects of the Move action is made false! But clearly this is not what we want from a theory of updates! Even if we adopt the device of "protecting" certain formulas and so preserve all the general laws in the example above, we will still get two minimal models-- one where switch 2 moves when you move switch 1, and the other where the light goes on. Moreover, we cannot protect the value of switch 2 in general, for there are some circumstances in which we may say that moving switch1 will change switch 2 (such as if a child is playing with the switches and trying to keep the light off as an adult switches switch 1).

We agree with Winslett (1990), Asher (1991) and others that a satisfactory treatment of persistence in beliefs and belief updating must use some nonmonotonic formalism, and this motivates

---

[2]The origin of the observation that Gärdenfors's notion of updating is inappropriate for updating under the assumption of change is due originally to Keller and Winslett (1985).

our approach. In what follows, we will not try to specify exactly what the nonmonotonic formalism is. We now set out in somewhat programmatic fashion our framework.

I. THE MONOTONIC FOUNDATION

We now present briefly the framework of our theory. First, we define models for the monotonic core attitudes, as well as for action and time. Our language L_{BD} is largely familiar from dynamic logic with quantification over restricted programs. There are terms for actions (or, more generally, events). Events or actions may combine to form complex events that we call programs. L_{BD} has two special operators, $\mathcal{D}$ and $\mathcal{R}$, that convert action terms into formulas: $\mathcal{D}(\pi)$ means that π is just now being initiated, and $\mathcal{R}(\pi)$ means that π has just been completed. Formulas containing $\mathcal{D}$ or $\mathcal{R}$ are activity formulas, and must be distinguished from state formulas, which lack these operators.

As regards state formulas, L_{BD} is an ordinary modal, propositional language with an operator K for belief, a nonmonotonic conditional >, and an historical necessity operator A. The conditional > can be integrated into a theory of nonmonotonic inference, such as the Commonsense Entailment system of Asher and Morreau (1991). Here is a list of the basic symbols of L_{BD}:

Propositional constants: p, q, r, ...
Basic action constants: α, β, γ,, ...
Program constants: μ_1, ...
Basic-action-to-program function symbol: $\delta(\alpha)$
Program-to-program function symbols:
$(\pi|\pi')$, $(\pi;\pi')$, π^*,
Proposition-to-program function symbol: $?\varphi$
Program-to-proposition operators: $\mathcal{R}(\pi)$, $\mathcal{D}(\pi)$
Proposition-to-proposition operators: $\varphi\&\psi$, $\neg\varphi$, $K\varphi$, $A\varphi$, $\varphi>\psi$, $N\varphi$, $L\varphi$

Semantics of L_{BD}

Our model theory makes use of *scenarios*. They can be thought of as worlds extended indefinitely into the future with a designated present moment, which we call the *root*. Several scenarios may share the same root. This samerootedness of scenarios is a primitive in our model and gives the effect of branching time.

L_{BD} frames are tuples $\langle BASIC,\$,\Pi,*,K,\approx\rangle$, where \$ is the set of scenarios, $BASIC \subseteq \wp(\$ \times \$)$ $\Pi \subseteq \wp(\$ \times \$)$, $\approx$ a relation in $\$\times\$$, K a binary relation on \$, and * a function from $\$\times\wp(\$) \rightarrow \wp(\$)$.

The belief accessibility relation K is transitive, and serial. Consequently, the logic for the belief operator is KD4.

We adopt as constraints on the models that the samerootedness relation, $\approx$, is an equivalence relation. In addition, we require the following constraints on Π

1. $BASIC \subseteq \Pi$
2. Π is closed under union, composition, and the following operation, for every subset A of \$: if π, π' $\in \Pi$, then $(\pi \restriction A) \cup (\pi' \restriction \$-A) \in \Pi$.

We can define a subscenario (temporal order) relation:

$$s \subset s' \text{ iff there is an } R \in \Pi\ \langle s,s'.\rangle \in R.$$

Finally, the selection function * must also satisfy two constraints:

Facticity: $*(s,A) \subseteq A$.
Dudley Doorite: $*(s,A\cup B) \subseteq *(s,A) \cup *(s,B)$

Fact 0 (Morreau 1992): Dudley Doorite + Facticity $\Rightarrow$ Specificity, viz. If $A \subseteq B$ & $*(s,A) \cap *(s,B) = \emptyset$, then $*(s,B) \cap A = \emptyset$.

This fact is very useful since it allows us to prove that more specific defaults always win over less specific ones in the event of a conflict.

A BD frame together with an interpretation function $[\![\ ]\!]$ yields a BD model M. The interpretation of an atomic state formula is a set of scenarios. Interpretations of program terms must satisfy the following constraints (from standard propositional dynamic logic):

1. For basic action term α, $[\![\alpha]\!] \in$ Basic, and $[\![\delta(\alpha)]\!] = [\![\alpha]\!]$.
2. For each proper program constant π, $[\![\pi]\!] \in \Pi$.
3. $[\![(\pi;\pi')]\!] = [\![\pi]\!] \circ [\![\pi']\!]$
4. $[\![(\pi \mid \pi')]\!] = [\![\pi]\!] \cup [\![\pi']\!]$
5. $[\![?\varphi]\!] = \{<s,s>: M,s \models \varphi\}$

The semantics for the operators is given by the following clauses:

$M,s \models K\varphi$ iff $\forall s' (sKs' \rightarrow s' \in [\![\varphi]\!]_{M,})$
$M,s \models A\varphi$ iff $\forall s'(s \approx s' \rightarrow M, s' \models \varphi)$
$M,s \models \mathcal{R}(\pi)$ iff $s \in \text{Range}(R_\pi)$
$M,s \models \mathcal{D}(\pi)$ iff $s \in \text{Domain}(R_\pi)$
$M,s \models N\varphi$ iff $\exists\alpha \in \text{BASIC}\ \exists s' (<s,s'> \in R_\alpha \ \&\ M,s' \models \varphi)$
$M, s \models L\varphi$ iff $\exists\alpha \in \text{BASIC}\ \exists s' (<s',s> \in R_\alpha \ \&\ M,s' \models \varphi)$

The operator K represents the modality of hard (unrevisable) belief, and the operator A gives us a notion of historical necessity. Finally, $\mathcal{D}$ and $\mathcal{R}$ are dual operators that yield the initial state of a program or its final state. Intuitively, $\mathcal{D}(\pi)$ means that the action π is being started, and $\mathcal{R}(\pi)$ means that π has just been completed. They help define the traditional temporal and dynamic logic operators. $N\varphi$ means that φ is true in the next moment, after some basic action has been completed, and $L\varphi$ means that φ was true in the last moment. Thus, the temporal structure is complete and closely tied to the execution of actions. The following definitions introduce the more familiar dynamic modalities $[\pi]$ and $<\pi>$, as well as abbreviations $F\pi\varphi$ and $P\pi\varphi$, representing φ's being true after the completion and immediately before the initiation of program π.

DEFINITIONS

$[\pi]\ \varphi$ iff $A(\mathcal{D}(\pi) \rightarrow \mathcal{D}(\pi; ?\varphi))$
$<\pi>\ \varphi$ iff $\neg[\pi]\neg\varphi$
$F\pi\ \varphi$ iff $\mathcal{D}(\pi;?\varphi)$
$P\pi\varphi$ iff $\mathcal{R}(?\varphi;\pi)$

The axiomatization of BD consists of familiar axioms and rules of temporal, dynamic, conditional and doxastic logics. [See Appendix A]

Theorem 1: BD has a canonical model and is complete for the specified class of frames.

Lemma 1.1: Any BD consistent set of formulas can be extended to a maximal consistent saturated set.

Lemma 1.2: BD contains all the relevant Barcan formulas. Hence the following sets are omega-complete, given that s is: $\{\varphi: A\varphi \in s\}$, $\{\varphi: F_\pi\ \varphi \in s\}$, $\{\varphi: K\varphi \in s\}$, $\{\varphi: \psi > \varphi \in s\}$.

Lemmas 1.1 and 1.2 allow us to build the canonical model for BD in the usual manner. Below we let π range over proper programs and α over atomic actions. $M_{CAN} = <\$_{CAN}, BASIC_{CAN}, \Pi_{CAN}, K_{CAN}, *_{CAN}, \approx_{CAN}, [\![\ ]\!]_{CAN}>$, where:

$\$_{CAN}$ = the set of maximal consistent saturated sets of sentences of LBD
$[\![\pi]\!]_{CAN} = \{<s,s'>: \mathcal{D}(\pi) \in s \ \& \ \{\varphi: F_{\pi}\ \varphi \in s\} \subseteq s'\}$, for π a program or action constant
$[\![\varphi]\!]_{CAN} = \{s: \varphi \in s\}$, for φ atomic formula
$BASIC_{CAN} = \{[\![\alpha]\!]_{CAN}: \alpha$ a basic action term$\}$
$\Pi_{CAN} = \{[\![\pi]\!]_{CAN}: \pi$ a program term $\}$
$Succ_{CAN} = \bigcup \{[\![\alpha]\!]_{CAN}: \alpha$ a basic action term$\}$ (This enables us to define a linear temporal ordering $\subset_{CAN}$ on $\$_{CAN}$.)
$K_{CAN} = \{<s,s'>: \{\varphi: K\varphi \in s\} \subseteq s'\}$
$*_{CAN}(s, [\![\varphi]\!]) = \{s': \{\psi: \varphi > \psi \in s\} \subseteq s'\}$
$\approx_{CAN} = \{<s,s'>: \{\varphi: A\varphi \in s\} \subseteq s'\}$

We now have just three more lemmas to prove completeness:
Lemma 1.3 : M_{CAN} is a model for BD.
Lemma 1.4 (Henkin Lemma): $\forall s \in \$_{CAN}\ M_{CAN}, s \models \varphi$ iff $\varphi \in s$
Lemma 1.5: In M_{CAN}, the temporal order is transitive, discrete and linear for any finite period, $\approx$ is an equivalence relation, and K is transitive and serial.

The first and third lemmas can be verified in the usual way. The proof of the second lemma has two non-trivial parts: the induction steps for the formulas $\mathcal{D}(\pi)$ and $\mathcal{R}(\pi)$. One must do an induction on the complexity of program terms.

II. NONMONOTONIC SUPERSTRUCTURE

As we mentioned in the introduction, we take a novel approach to the question of belief revision by distinguishing between hard (incorrigible) and soft (revisable) belief. Moreover, we use a theory of nonmonotonic reasoning to provide a principled account of revision. The soft beliefs are simply the nonmonotonic consequences of the hard beliefs. The soft intentions depend upon soft beliefs and our hard commitments.

We describe the nonmonotonic component of our theory abstractly as an operation on sets of scenarios. Intuitively, NORM(X) is a subset of X that captures the information that is nonmonotonically inferable from TH(X), where X represents the set of doxastic alternatives. We will build on the theory of Commonsense Entailment developed by Asher and Morreau (1991). We construct three alternative accounts of nonmonotonic consequence, symbolized $|\!\approx$. All three share certain crucial features, which will be used in developing our account of belief revision and update. These features are:

- All three license a defeasible version of modus ponens.
- All three incorporate the principle of Specificity (defaults with more specific conditions override conflicting defaults with less specificity). In addition, in all three theories, this feature is a consequence of the underlying monotonic conditional logic and is not secured by any ad hoc device, such as prioritizing defaults.
- All three permit arbitrary nesting of defaults.
- All three are resilient when irrelevant information is added to the premises.

The presence of these features provides our account of belief revision and update with a number of very desirable features. First of all, the revisions and updates will be fully determined by the object-level sentences used to represent the agent's beliefs. There is no need to bring in additional machinery, such as degrees of entrenchment. Secondly, when revision and update are applied to extremely complex theories, the Specificity feature of the nonmonotonic logic takes care of the correct prioritizing of defaults. In such cases, alternative approaches, like prioritized circumscription, impose the non-trivial task of formulating the correct priorities in a case-by-case fashion. Finally, when dealing with multi-agent problems, the ability to nest defaults is very desirable.

The first version of nonmonotonic inference, $NORM_1$, is simply defined in terms of the original version of Commonsense Entailment:

$$NORM_1(X) = \{s: \forall\varphi(TH_M(X) \models\!\approx_1 \varphi \rightarrow M,s \models \varphi)\}$$

In Asher and Morreau (1991), an inductive definition of $\models\!\approx_1$ is given in terms of the canonical model for the underlying monotonic logic.

Definition: Let σ, $p \subseteq W_{CAN}$, the set of worlds in the canonical model.

$NORM(\sigma, p) = \{w \in \sigma: w \notin p \setminus *(\sigma, p)\}$, if $\sigma \cap *(\sigma, p) \neq \emptyset$
$= \sigma$ otherwise.

Now let $ANT(\Gamma) = \{\psi: \psi > \varphi$ is a subformula of a formula in $\Gamma\}$.

Definition: $\Gamma(\zeta)$ the Γ-Normalization sequence for a given ordering ζ on $ANT(\Gamma)$

$\Gamma_0(\zeta) = W_{CAN} \cap \| \Gamma \|$

$\Gamma_{\beta+1}(\zeta) = NORM(\Gamma_\beta(\zeta), p)$, for $\zeta(p) = n+1$ & $\beta + 1 = \lambda + km + n +1$, where m is the size of $ANT(\Gamma)$, λ is limit ordinal, and k is a natural number.

$\Gamma_\lambda(\zeta) = \cap_{\beta < \lambda} \Gamma_\beta(\zeta)$

Definition: $\mathcal{B}$ is a Γ-fixpoint for a Γ normalization sequence $\Gamma(\zeta)$ iff $\exists\beta$ such that $\Gamma_\beta(\zeta) = \mathcal{B}$ & $\forall\alpha \geq \beta$ $\Gamma_\beta(\zeta) = \Gamma_\alpha(\zeta)$.

Fact 2: For every ordering ζ of $ANT(\Gamma)$ and Γ-normalization sequence $\Gamma(\zeta)$, there is a Γ fixpoint.

Definition: $\Gamma \models\!\approx \varphi$ iff for every Γ fixpoint $\mathcal{B}$, $\mathcal{B} \models \varphi$.

Our second version of normalization, $NORM_2$, is defined directly on a given model M. We define an M-normalization function f to be a function from $\wp(\$) \rightarrow \wp(\$)$ such that

$f(X) \subseteq X$
$f(Y) \subseteq X$ & $f(X) \subseteq Y \Rightarrow f(X) = f(Y)$
$X \neq \emptyset \Rightarrow f(X) \neq \emptyset$

We now define the *score* of a M-normalization function f such that :

$$\langle A, B\rangle \in Score(f) \text{ iff } [f(A) \subseteq B \rightarrow f(A) \subseteq *(f(A), B)],$$

and we say that f is a *maximal M*-normalization function iff there is no g such that score(f) $\subset$ score(g). Now we define:

$$NORM_2(X) = \cup\{f(X): f \text{ is a maximal M-normalization function}\}$$

Fact 3: $NORM_2$ is an M-normalization function.

The third version of normalization, $NORM_3$, corresponds closely to a new version of commonsense entailment, developed independently by Michael Morreau and by Andrew Schwartz (a graduate student in philosophy at the University of Texas) and Koons. First, we impose an ordering on scenarios, preferring those scenarios that minimize anomalies (exceptions to defaults).

$$ANOM_3(s) = \{p \in \wp(\$): s \in p \setminus *(s,p) \}$$

$$s << s' \text{ iff } ANOM_3(s') \subset ANOM_3(s)$$

$NORM_3(X)$ is simply the set of <<-maximal elements of X:

$$NORM_3(X) = \{s \in X: \neg\exists s'(s' \in X \& s << s')\}$$

There are metalogical advantages and disadvantages to each of these definitions. The following are some well-known metalogical properties:

(ia) $\Gamma, \varphi \mid\approx \psi \;\&\; \vDash (\varphi \leftrightarrow \alpha) \Rightarrow \Gamma, \alpha \mid\approx \psi$	(LE)
(ib) $\Gamma \vDash \varphi \Rightarrow \Gamma \mid\approx \varphi$	(Supraclassicality)
(ii) $\Gamma \mid\approx \varphi \;\&\; \Gamma, \varphi \mid\approx \psi \Rightarrow \Gamma \mid\approx \psi$	(Cut)
(iii) $\Gamma \mid\approx \varphi \;\&\; \Gamma \mid\approx \psi \Rightarrow \Gamma, \varphi \mid\approx \psi$	(C-Mon)
(iv) $\Gamma, \alpha \mid\approx \varphi \;\&\; \Gamma, \beta \mid\approx \varphi \Rightarrow \Gamma, (\alpha \vee \beta) \mid\approx \varphi$	(OR)
(v) $\Gamma \mid\approx \psi \;\&\; \text{not}(\Gamma \mid\approx \neg\varphi) \Rightarrow \Gamma, \varphi \mid\approx \psi$	(R-Mon)

Consequence relations having properties (i) -- (iv) are called "preferential", while those with all five properties are called "rational" (Lehmann 1989). The corresponding properties of the normalization function would be these:

(i) $NORM(X) \subseteq X$
(ii) If $NORM(X) \subseteq Y$, then $NORM(X) \subseteq NORM(X \cap Y)$
(iii) If $NORM(X) \subseteq Y$, then $NORM(X \cap Y) \subseteq NORM(X)$
(iv) If $NORM(X) \subseteq Z$ & $NORM(Y) \subseteq Z$, then $NORM(X \cup Y) \subseteq Z$
(v) If $NORM(X) \cap Y \neq \emptyset$, then $NORM(X \cap Y) \subseteq NORM(X)$

Fact 4:
(a) $NORM_1$ has properties (i) and (ii) only.
(b) $NORM_2$ has properties (i), (ii) and (iii), but not (iv) or (v).
(c) $NORM_3$ has properties (i) through (iv), but not (v).

If we restrict α and β to be > free formulae and define $\mid\approx$ relative to a fixed background theory of unnested rules T in the unnested fragment of our language L_{CE}, then we can show that $\mid\approx_1$ is a rational consequence relation.

Fact 5: Let α, β, γ be > free formulae and let T be a set of simple unnested > formulae or strict conditionals.

(ia) $T, \alpha \vDash \beta \Rightarrow T, \alpha \mid\approx \beta$	SUPRACLASSICALITY
(ib) $T, \alpha \mid\approx \beta \;\&\; \vdash \alpha \leftrightarrow \gamma \Rightarrow T, \gamma \mid\approx \beta$	LE
(iii) $T, \alpha \mid\approx \beta \;\&\; T, \alpha, \beta \mid\approx \gamma \Rightarrow T, \alpha \mid\approx \gamma$	CUT
(iv) $T, \alpha \mid\approx \beta \;\&\; T, \alpha \mid\approx \gamma \Rightarrow T, \alpha, \beta \mid\approx \gamma$	CM
(v) $T, \alpha \mid\approx \beta \;\&\; T, \gamma \mid\approx \beta \Rightarrow T, \alpha \vee \gamma \mid\approx \beta$	OR
(vi) $T, \alpha \mid\approx \beta \;\&\; \text{not}(T, \alpha \mid\approx \neg\gamma) \Rightarrow T, \alpha \;\&\; \gamma \mid\approx \beta$	RM

Proof: See Appendix B.

Moreover, since $\mid\approx_1$ is defined inductively, it possesses certain proof-theoretic advantages. It is possible to define a 2-place function $\sqsupset$ that encodes the consequence relation $\mid\approx_1$, in the following sense.

Fact 6: $\varphi \mid\approx_1 \psi$ iff for all orderings ζ of the antecedents of subformulae of φ, $\{\varphi, \sqsupset(\varphi,\zeta)\} \vDash \psi$.
Proof: See Appendix C.

The function $\sqsupset$ is computable in the propositional case (or any case in which consistency is decidable). In the case of $|\!\approx_3$, it is known that there is no corresponding $\sqsupset$ function. At present, the corresponding question with respect to $|\!\approx_2$ is open.

If nesting of conditionals is limited, we believe that the $|\!\approx_3$ consequences of a finite set are r.e., for the propositional language L_{BD}. Let's say that a formula has a maximum nesting depth of 0 if no >-conditionals occur within the scope of > and a maximum depth of 1 if no conditionals of depth 1 occur within the scope of >. A set of formulas Γ has a maximum depth of 2 if every member has the property. Let Ant(Γ) be the set of antecedents of >-conditionals (of any depth) in Γ.

DEFINITION. $s \ll_{Ant(\Gamma)} s'$ iff
$\{\|\varphi\|: \varphi \in Ant(\Gamma) \ \&\ [s \in \|\varphi\| \rightarrow s \in *(s,\|\varphi\|)]\} \subset \{\|\varphi\|: \varphi \in Ant(\Gamma) \ \&\ [s' \in \|\varphi\| \rightarrow s' \in *(s',\|\varphi\|)]\}$

Claim: Suppose Γ is a finite set of formulas of maximum depth 2. Suppose $\top \in Ant(\Gamma)$, and that there is a $\varphi \in Ant(\Gamma)$ such that for every ψ, if $\Gamma \models \varphi > \psi$, then $\Gamma \models \psi$. A scenario s is a $\ll$-maximal member of $\|\Gamma\|$ in M_{CAN} iff s is a $\ll_{Ant(\Gamma)}$-maximal member of $\|\Gamma\|$ in M_{CAN}.

This claim means that, in verifying that a formula φ is among the $|\!\approx_3$ consequences of a finite set Γ, one has to show only that every $\ll_{Ant(\Gamma)}$-maximal member of $\|\Gamma\|$ in M_{CAN} verifies φ, that is, one has to show that φ is verified in any scenario that minimizes the anomalies among the set of antecedents of Γ.[3]

In the rest of the paper, whenever we use the expression "NORM" without specifying any one of the three consequence relations, what is said is true however NORM is specified.

Armed with a nonmonotonic consequence relation, we now extend the language to include operators for soft beliefs . We use NORM to characterize the soft beliefs by defining basic information and soft commitment states.

DEFINITION <u>BEL</u>(s) = NORM({s': sKs'})

III. DYNAMICS OF BELIEF

The point of our framework is to be able to explore belief revision in a nonmonotonic setting We are now able to define these notions.

DEFINITION: Rev(s, $\|\varphi\|$) = NORM({s': sKs'} $\cap$ $\|\varphi\|$)

Belief revision is defined simply as the result of applying the CE nonmonotonic logic to the new body of (hard) information. There is no need for a ghost in the machine to tease out the right revisions in a case-by-case fashion. The agent's dispositions to learn and revise are expressed declaratively as objects of the agent's belief. Moreover, the language can be extended with new modal operators, without disturbing the definitions of Rev or NORM.

The definition of belief revision above has a number of desirable formal properties.

Fact 7: (i) Rev(s, $\|\varphi\|$) $\models \varphi$ $\cap K(s)$ (Success)
(ii) Rev(s, $\|\varphi\|$) = $\emptyset$, only if $\|\varphi\| = \emptyset$ (Consistency)
(iii) Rev(s, $\|\varphi\|$) $\subseteq$ {s': sKs' } (Elimination)
(iv) NORM(Rev(s,$\|\varphi\|$)) = Rev(s,$\|\varphi\|$) (Closure)

Fact 8: If <u>BEL</u>(s) $\subseteq \|\varphi\|$, then Rev(s,$\|\varphi\|$) = <u>BEL</u>(s).

[3] We haven't been able to verify what happens in the case of $NORM_3$ when the depth of nesting is greater than 2.

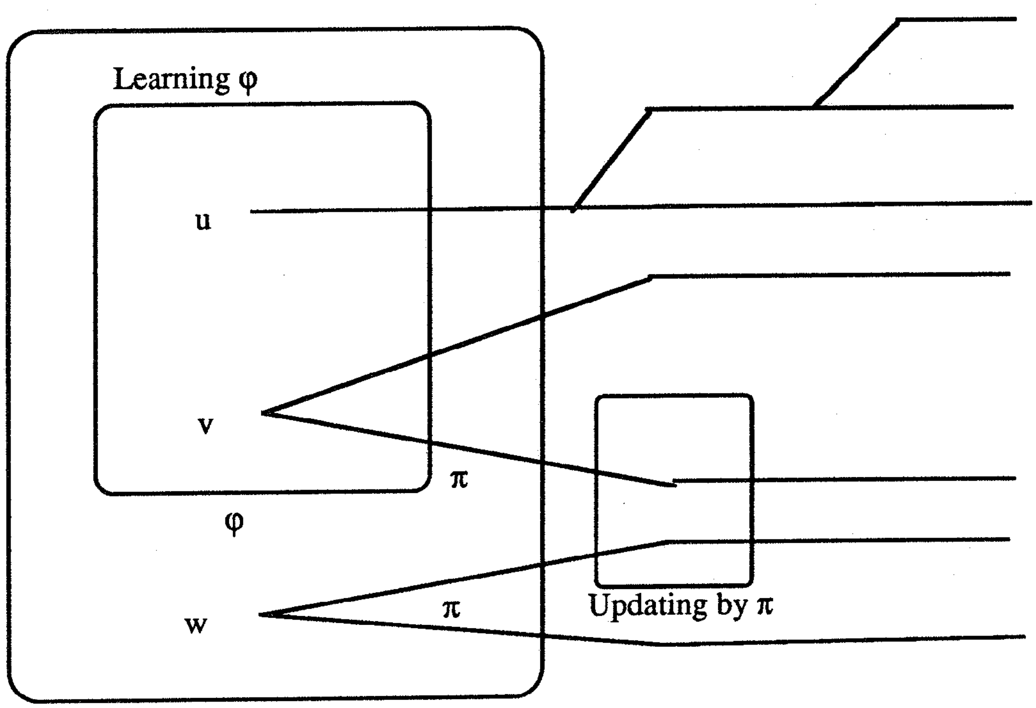
Fact 9: If $BEL_3(s) \cap \|\varphi\| \neq \emptyset$, then $BEL_3(s) \cap \|\varphi\| \subseteq Rev_3(s,\|\varphi\|)$.

Because our language has dynamic operators and the imaging operators $\mathcal{D}$ and $\mathcal{R}$, we can define a uniform theory of updating and revising beliefs in a changing world. Updating and non-update revising are in fact quite close. Revision without update keeps the current state fixed but adds information about past or future states. Updating with information about an observed change always involves formulae in the form $\mathcal{R}(\alpha)$ and necessarily shifts the location of the doxastic "now". Incorporating this change into the agent's beliefs involves two steps. First, the agent must calculate what the expansion or revision of his current belief state with $\mathcal{D}(\alpha)$ would yield. Secondly, he must adjust what he takes to be the current state of the world to a new temporal slice, the time after the observed event α. This second effect corresponds to the imaging operation of Katsuno and Mendelzon.

DEFINITION: $\text{Update}(s,\alpha) = \{s': \exists s'' \langle s'',s' \rangle \in R_\pi \ \& \ s'' \in Rev(s, \|\mathcal{D}(\alpha)\|)$

Here is a schematic picture of how observations of events and the learning of static information about states affect a belief state in different ways. Suppose that σ contains the following three scenarios u, v, and w.

The larger rectangle on the left represents the initial belief state σ of the agent. The smaller rectangle within σ represents the proposition $\|\varphi\|$ (or, perhaps, the intersection of σ with $\|\varphi\|$). If the agent learns φ (without updating), then the corresponding case of the resulting belief state will be represented by this smaller rectangle. Alternatively, if the agent updates by observing event or action π (without learning anything, in our sense), then the resulting case of the new belief-state is just the image of σ under the relation R_π (represented by the rectangle on the right). Of course, in many cases, both learning and updating will take place. (This last possibility is not pictured.)



At this point, we would like to introduce two new kinds of basic actions or events: learn(φ) and observe(π). The basic action term 'learn(φ)' refers to an event in which the agent expands or revises his belief-state with the new information corresponding to the sentence φ. The term 'observe(π)' is a term designating an update event in which the agent observes that the event π has just occurred. A BD model in which there is a correspondence between the occurrence of events of learning and observing and the actual changes in the agent's belief-states is called a "doxastically faithful model". In each state, there are four possibilities: the agent both observes a change and learns new information, (ii) the agent observes a change but learns nothing new, (iii) the agent learns new information without observing any change, and (iv) the agent's belief-state is unchanged. If state s is in the domain of some event of the form learn(φ), we will say that s $\in$ LEARN, and if s is in the domain of some event of the form observe(π), we will say that s $\in$ OBSERVE.

DEFINITION. M is a *doxastically faithful model* iff for every $s \in \$_M$:
(1) If $s \in \mathcal{D}(\text{learn}(\varphi))$ & $s \in \mathcal{D}(\text{observe}(\pi))$, then $\underline{\text{BEL}}(\text{succ}(s)) = \text{Update}(\text{Rev}(s,\|\varphi\|), \pi)$.
(2) If $s \in \mathcal{D}(\text{learn}(\varphi))$ & $s \notin$ OBSERVE, then $\underline{\text{BEL}}(\text{succ}(s)) = \text{Rev}(s,\|\varphi\|)$.
(3) If $s \in \mathcal{D}(\text{observe}(\pi))$ & $s \notin$ LEARN, then $\underline{\text{BEL}}(\text{succ}(s)) = \text{Update}(s,\pi)$.
(4) If $s \notin$ LEARN $\cup$ OBSERVE, then $\underline{\text{BEL}}(\text{succ}(s)) = \underline{\text{BEL}}(s)$.

Persistence and the Causation of Change

We now have the resources for a general theory of the presumption of persistence and the form of reasoning about the causation of change. Firstly, we will need a special binary modal operator $\boxdot$ to represent a relation of evidential dependency. The formula $\boxdot(\varphi,\{\psi_1,\ldots,\psi_n\})$ will express a conceptual dependency of φ on the members of the set $\{\psi_1,\ldots,\psi_n\}$. Intuitively, this means that φ is essentially disjunctive in character, and each ψ_i represents an agglomeration of atomic facts holding in one of the ways in which φ could be true.

A model will include a conceptual dependency relation C, which is to be a function from $\$ \times \wp(\$)$ to finite subsets of $\wp(\$)$.

$$M,s \models \boxdot(\varphi,\{\psi_1,\ldots,\psi_n\}) \text{ iff } C(s,\|\varphi\|) = \{\|\psi_1\|,\ldots,\|\psi_n\|\}$$

The following axiom schemata are valid:

$\vdash \boxdot(\varphi,\{\psi_1,\ldots,\psi_n\}) \rightarrow \boxdot(\varphi,\{\zeta_1,\ldots,\zeta_n\})$, where $\{"\psi_1",\ldots,"\psi_n"\} = \{"\zeta_1",\ldots,"\zeta_n"\}$
$\vdash \boxdot(\varphi,\{\psi_1,\ldots,\psi_n\}) \rightarrow \boxdot(\zeta,\{\psi_1,\ldots,\psi_n\})$, where $\vdash \varphi \leftrightarrow \zeta$
$\vdash \boxdot(\varphi,\{\psi_1,\ldots,\psi_n\}) \rightarrow \boxdot(\varphi,\{\psi_1,\ldots,\psi_{j-1},\zeta,\psi_{j+1},\ldots,\psi_n\})$, where $\vdash \psi_j \leftrightarrow \zeta$
$\vdash \boxdot(\varphi,S)$ & $\boxdot(\varphi,S') \rightarrow \boxdot(\varphi,S'')$, where $S'' = S \cup S'$

Suitable restrictions on C will validate the following schema:

$\boxdot(\varphi,\{\psi_1,\ldots,\psi_n\}) \rightarrow \Box(\varphi \leftrightarrow (\psi_1 \vee \ldots \vee \psi_n))$
$\boxdot(\varphi,\{\psi_1,\ldots,\psi_n\})$ & $\Diamond\varphi \rightarrow \Diamond\psi_i$

If the language we are using happens to mirror the conceptual scheme of every agent at every scenario, we could add the following axioms:

$\boxdot(\varphi,\{\varphi\})$, where φ is a conjunction of literals
$\boxdot((\varphi_1 \vee \ldots \vee \varphi_n), \{\varphi_1,\ldots,\varphi_n\})$, where each φ_i is a conjunction of literals, and the disjunction $(\varphi_1 \vee \ldots \vee \varphi_n)$ is logically equivalent to no shorter formula

However, this is not an assumption that we would wish to make generally. We want to be able to reason about agents with conceptual schemes radically different from our own, or with schemes that vary from state to state. Moreover, we often find it convenient to work with a language that does not reflect our own conceptual scheme, introducing, for example, abbreviations for logically complex propositions. It is methodologically crucial to represent explicitly those features of the agent that determine the correct nonmonotonic inference, and not to rely on linguistic form as an implicit representation of these features.

Next, we introduce three additional axioms: the law of inertia, the law of the priority of causation over inertia, and the law of frame dependency.

LAW OF INERTIA

$$\psi > (\psi > N\psi)$$

PRIORITY OF CAUSATION OVER INERTIA

$$(\varphi > N\neg\psi) \rightarrow [\, (\psi \,\&\, \varphi) > \neg(\psi > N\psi)\,]$$

FRAMEWORK DEPENDENCE

$$[\boxdot(\psi,\{\ldots\varphi\ldots\}) \,\&\, \neg(\varphi > N\varphi)] \rightarrow \neg(\psi > N\psi)$$

Note especially that in the law of causal priority and the law of Framework Dependence, the main connective is the truth-functional conditional, and not >. Let BD+ be the logic BD extended by the three axiom schemata above, as well as by an S5 logic for $\Box$ and the schemata introduced above for $\boxdot$.

EXAMPLE #1: THE AUSTIN MERCADO

Let's apply this structure to the two examples in the first section of the paper. For Example #1, we will use the following abbreviations:

α: tomatillos in the basket
β: chiles in the basket
γ: $\alpha \leftrightarrow \neg\beta$
π: the action of adding tomatillos to the basket

We will assume that γ is evidentially dependent on the constituents of $(\alpha \,\&\, \neg\beta)$ and $(\neg\alpha \,\&\, \beta)$.

$$\boxdot(\gamma, \{(\alpha \,\&\, \neg\beta), (\neg\alpha \,\&\, \beta)\})$$

We will also assume that the agent believes a defeasible causal law of the form:

$$(L_1)\quad \mathcal{D}(\pi) > N\alpha$$

[Remember: the operator 'N' represents "in the next moment..."] Finally, we will assume that the agent observes π, and that the model we're using is doxastically faithful. Let s designate the state before the observed action and s' the succeeding state. We'll assume that the theory of {s': sKs'} is the set of BD+-consequences of the set $\{\gamma, (L_1)\}$. By the law of the priority of causation, we know that :

$$\mathrm{Rev}(s,\mathcal{D}(\pi)) \models (\neg\alpha \,\&\, \mathcal{D}(\pi)) > \neg(\neg\alpha > N\neg\alpha)$$

Since the antecedent of this conditional is more specific than $\neg\alpha$, the antecedent of the relevant instance of the law of inertia, we have:

$$\mathrm{Rev}(s,\mathcal{D}(\pi)) \models \neg(\neg\alpha > N\neg\alpha).$$

A similar application of the priority of causation yields:

$$\mathrm{Rev}(s,\mathcal{D}(\pi)) \models \neg((\neg\alpha \,\&\, \beta) > N(\neg\alpha \,\&\, \beta)).$$

Finally, given the dependency of γ on $(\neg\alpha \ \& \ \beta)$, an application of the framework dependency principle yields the result:

$$\mathrm{Rev}(s,\mathcal{D}(\pi)) \models \neg(\gamma > N\gamma).$$

Since the agent cannot apply inertia to γ, the agent comes to believe α, but does not continue to accept γ or embrace $\neg\beta$. So, $\mathrm{Update}(s,\pi) \models \alpha$, but not $\mathrm{Update}(s,\pi) \models \neg\beta$. Since M is doxastically faithful, $\underline{\mathrm{BEL}}(s') \models \alpha$, but not $\underline{\mathrm{BEL}}(s') \models \neg\beta$.

EXAMPLE #2: THE LIFSCHITZ LAMP

For example #2, we will again use abbreviations.

α: switch 1 up
β: switch 2 up
γ: light on
π: flip switch 1 down

We assume that the agent believes, in the initial state s, $\alpha \ \& \ \neg\beta \ \& \neg\gamma$. In this case, we'll assume that the agent accepts a defeasible causal law of the form:

$$(L_2) \ \mathcal{D}(\pi) > N\neg\alpha$$

Moreover, we'll assume that the agent accepts a law of the form:

$$(L_3) \ N(\alpha \leftrightarrow \beta) > N\gamma$$

Finally, we'll assume that the agent observes π, resulting in state s'.

As before, the principle of the priority of causation, together with the specificity preference of the underlying nonmonotonic logic, guarantees that the agent believes in the non-applicability of inertia to α. However, we can construct two kinds of fixed-points: those in whch $\neg\gamma$ is abnormal because $N\gamma$ holds, and those in which $\mathcal{D}(\pi)$ is abnormal because $N\neg\alpha$ fails to hold. The description of the case is too weak to enable us to ensure that the tendency of the light to persist might cause the action π to fail to have its normal effect on switch1. If we replaced (L_2) by strict causal law, like $[\pi]N\neg\alpha$, the problem disappears and our nonmonotonic logic immediately gives the correct answer. Alternatively, we could replace (L_2) by $(L_2{*})$, which explicitly states that the light's being off has no effect on π's influence on the switch:

$$(L_2{*}) \ \mathcal{D}(\pi) \ \& \ \neg\gamma > N\neg\alpha$$

In either case, we are in a position to draw the conclusion that π has its normal effect.

$$\mathrm{Rev}(s,\mathcal{D}(\pi)) \models \neg N\alpha$$

The causal law L_3, together with the law of causal priority, monotonically entails:

$$(N(\alpha \leftrightarrow \beta) \ \& \ \neg\gamma) > \neg(\neg\gamma > N\neg\gamma)$$

Once again, the underlying conditional logic guarantees that this formula, together with the relevant instance of inertia, viz. $\neg\gamma > (\neg\gamma > N\neg\gamma)$, entails: $\neg\gamma > \neg[N\alpha \leftrightarrow N\beta]$. Let's assume that this biconditional is conceptually disjunctive:

$$\boxdot(\neg(N\alpha \leftrightarrow N\beta), \{(N\alpha\&\neg N\beta),(N\alpha\&\neg N\beta),\})$$

Since Rev(s,$e\mathcal{D}(\pi)$) supports $\neg N\alpha$, contradicting the first of the two constituent cases, normalization of $\neg\gamma$ is blocked, regardless of what happens to the value of $N\beta$. Consequently, we can apply inertia to $\neg\beta$ without introducing any new abnormality. So,

$$\mathrm{Rev}(s,\mathcal{D}(\pi)) \models N\neg\beta$$

Consequently,

$$\mathrm{Rev}(s,\mathcal{D}(\pi)) \models N(\alpha \leftrightarrow \beta).$$

Another application of the priority of causation (involving L_3) gives us:

$$\mathrm{Rev}(s,\mathcal{D}(\pi)) \models \neg(\neg\gamma > N\neg\gamma).$$

Finally, the agent's belief in L_3 entails that

$$\mathrm{Rev}(s,\mathcal{D}(\pi)) \models N\gamma.$$

Since the model is doxastically faithful, BEL(s') $\models \gamma \,\&\, \neg\beta$; i.e., the light goes on, and switch 2 does not move.

EXAMPLE #3: THE YALE SHOOTING PROBLEM

We'll abbreviate the statement of the problem as follows:

α: alive
β: loaded
π_0: wait
π_1: shoot

The relevant causal law schema is:

$$(L'')\ N^n(\mathcal{D}(\pi_1) \,\&\, \beta) > N^{n+1}\neg\alpha$$

Informally, if, in n moments from now, β holds and π_1 is done, then normally $\neg\alpha$ would hold in the n+1st moment. To deal with this case, we must generalize the law of the priority of causation to apply to arbitrary distances in the future.

GENERAL LAW OF CAUSAL PRIORITY

$$(N^n\varphi > N^{n+1}\neg\psi) \,\&\, N^n\varphi > \neg(N^n\psi > N^{n+1}\psi)$$

Unfortunately, this version of causal priority lands us in exactly the predicament suffered by circumscription (which is perhaps not surprising, since Commonsense Entailment is a descendant of circumscription). We are forced to choose between introducing an anomaly involving $N\alpha$ and one involving β. We can avoid producing a counterexample to inertia as applied to $N\alpha$ by producing a counterexample to inertia as applied to β, i.e., by supposing that the gun mysteriously unloads itself while we wait.

The solution involves a modification of our definition of nonmonotonic entailment that incorporates the information encoded by the $\boxdot$ modality. In the case of $NORM_1$, we must revise the definition of the normalization function. When normalizing the antecedent ψ at stage β, it will no longer be sufficient to check that every φ such that $\sigma \models \psi > \varphi$ is consistent with σ; instead, one must check that every proposition θ upon which some such φ is dependent is consistent with σ. If normalizing with ψ would introduce an essentially disjunctive condition, and one's current information contradicts one of the constituent disjuncts, then normalization aborts and $N(\sigma,\psi) = \sigma$.

Definition: Let σ, $p \subseteq W_{CAN}$, the set of worlds in the canonical model.

$$\mathrm{NORM}(\sigma, p) = \{w \in \sigma : w \notin p \setminus *(\sigma, p)\},$$
$$\text{if for every } q \text{ such that } q \in C(c\sigma, *(\sigma,p)),\ \sigma \cap q \neq \emptyset$$
$$= \sigma \text{ otherwise.}$$

Returning to the Yale Shooting Problem, we find that the conditional logic incorporated into BD+ gives us the following result.

$$[\, N(\mathcal{I}(\pi_1) \,\&\, \beta) > NN\neg\alpha] \models [N\alpha > (N\neg\mathcal{I}(\pi_1) \vee N\neg\beta)]$$

The left-hand side is the relevant instance of our causal law, L". Our conditional logic guarantees that L", in the presence of the law of inertia, monotonically entails the right-hand side. This means that normalizing $N\alpha$ in our example introduces a disjunctive condition. Let's assume that this disjunction is conceptually dependent on its disjuncts.

$$\boxdot((N\neg\mathcal{I}(\pi_1) \vee N\neg\beta), \{N\neg\mathcal{I}(\pi_1), N\neg\beta\})$$

Since our information state contradicts, $N\neg\mathcal{I}(\pi_1)$, we cannot normalize $N\alpha$, even if our information state includes the information $N\neg\beta$ (that the gun has become unloaded). Thus, introducing an abnormality with respect to β does not enable us to eliminate an abnormality with respect to $N\alpha$. Consequently, every fixed point will contain $N\beta$ and $NN\neg\alpha$, the desired conclusions.

In addition, this modification of the nonmonotonic logic renders the Law of Frame Dependency redundant. Frame dependency is subsumed as a special case of the sensitivity of nonmonotonic inference to constitutive cases. Even in the absence of an explicit law of frame dependency, applications of inertia to conceptually disjunctive conditions will be blocked whenever there is change in any of the constituent disjuncts.

BIBLIOGRAPHY

Asher, N. (1991): "Belief Dynamics and DRT," in Fuhrmann A. & Morreau M. (eds.) *The Logic of Change*, Springer Verlag, 1991, pp. 282-321.

Asher, N. (1993): "Extensions for Commonsense Entailment", IJCAI Workshop.

Asher, N. & R. Koons (1993): "The Revision of Beliefs and Intentions in a Changing World, Working Notes AAAI Spring Symposium Committee, March 1993, Stanford, CA.

Asher, N. and M. Morreau (1991): "Commonsense Entailment: A Modal Theory of Nonmonotonic Reasoning", *IJCAI 91*, Sydney Australia, Morgan Kaufmann, San Mateo, Calif.

Asher, N. & M. Morreau (199+): "What Some Generics Mean," in Carlson, G., M. Krifka and F.J. Pelletier, eds., The *Generic Book*, Chicago, IL: Chicago University Press, in press.

Benferhat, S., D. Dubois & H. Prade (1992): "Representing Default Rules in Possibilistic Logic," *KR 92*, Cambridge MA.

Benferhat, S., C. Cayrol, D. Dubois, J. Lang & H. Prade (1993): "Inconsistency Management and prioritized syntax-based Entailment," *IJCAI 93*, Chambery.

Boutilier, C. (1990): "Conditional Logics of Normality as Modal Systems", AAAI.

Boutilier, C. (1992): *Conditional Logics for Default Reasoning and Belief Revision*, Ph.D. Thesis, University of British Columbia, Vancouver, BC, Canada.

Brewka, G. (1989): "Preferred Subtheories: an Extended Logical Framework for Default Reasoning," *IJCAI 89*.

Delgrande, J. (1988): "An Approach to Default Reasoning based on First Order Conditional Logic: Revised Report" *Artificial Intelligence* **36**, pp. 63-90.

Gärdenfors, P. (1988): *Knowledge in Flux*, MIT Press, Cambridge, Mass.
Goldszmidt, M. & J. Pearl (1990): "On the Relation between Rational Closure and System Z," *Third International Workshop on Nonmonotonic Reasoning*, Lake Tahoe, CA.
Kraus, S. D. Lehmann and M. Magidor (1990): "Nonmonotonic Reasoning, Preferential Models and Cumulative Logics," *Artificial Intelligence* **44**, pp. 167-207.
Lamarre, P. (1992): *Etude des Raisonnements Nonmonotones: Apports des Logiques des Conditionnels et des Logiques Modales*, Thèse de troisième cycle, Université Paul Sabatier, Toulouse.
Lehmann, D. (1989): 'What does a Conditional Knowledge Base Entail?', *KR 89*. Toronto, Canada.
Lehmann, D. (1992): "Another Perspective on Default Reasoning," Report TR-92-12, Institute of Computer Science, Hebrew University.
Lifschitz, V. (1987): ' Formal Theories of Action', *The Frame Problem in Artificial Intelligence*, Frank M. Brown (ed.), pp. 35-57. Morgan Kaufmann, Los Altos, Calif.
Lifschitz, V. (1990): ' Frames in the Space of Situations', *Artificial Intelligence* **40**.
Lifschitz, V. and A. Rabinov (1989): 'Miracles in formal theories of action', *Artificial Intelligence* **38**: 225-237.
Katsuno, H. and A. Mendelzon (1989): 'A Unified View of Propositional Knowledge Base Updates', *Proceedings of the 11th International Joint Conference on Artificial Intelligence*, Morgan Kaufmann, San Mateo, Calif., pp. 1413-1419.
Katsuno, H., and A. Mendelzon (1991): 'On the Difference Between Updating a Knowledge Base and Revising It', *Proceedings of the Second International Conference on Principles of Knowledge Representation and Reasoning*, Morgan Kaufmann Press, 387-394.
Keller, A. and M. Winslett (1985): 'On the Use of an Extended Relational Model to Handle Changing Incomplete Information', *IEEE Transactions on Software Engineering*, SE-11: 7, 620-633.
Manna, Z. and R. Waldinger (1987): 'How to Clear a Block: A Theory of Plans', *Journal of Automated Reasoning* **3**:343-378.
Moore, R. (1980): *Reasoning about Knowledge and Action*, Ph.D. Thesis, published as SRI Technical note 191, October 1980, SRI International.
Morreau, M. and H. Rott (1991): 'Is it Impossible to Keep Up to Date?', in *Nonmonotonic and Inductive Logic: First International Workshop*, J. Dix, K. Jantke, and P. Schmitt (eds.), Springer Verlag, Berlin, pp. 233-243.
Morreau, M. (1992): *Conditionals in philosophy and artificial intelligence*, Ph.D. thesis,Universität Stuttgart.
Pearl, J. (1988): *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann, San Mateo, Calif.
Pearl, J. (1990): "System Z: A Natural Ordering of Defaults with Tractable Applications to Nonmonotonic Reasoning," *TARK III*.
Pratt, V. (1980): 'Application of Modal Logic to Programming', *Studia Logica* **39**:257 - 274.
Rao, A. and N. Foo (1989): 'Minimal Change and Maximal Coherence: A Basis for Belief Revision and Reasoning about Action', *Proceedings of the 11th International Joint Conference on Artificial Intelligence*, Morgan Kaufmann, Los Altos, Calif., pp. 966-971.
Shoham, Y. (1988): *Reasoning About Change*, MIT Press, Cambridge, Mass.
Winslett, M. (1988): 'Reasoning about Action using a Possible Models Approach', in *AAAI-88: Proceedings of the Seventh National Conference on Artificial Intelligence*, pp. 89-93.
Winslett, M. (1990): 'Sometimes Updates are Circumscription', *AAAI* - 90.

Appendix A. Axiomatization of Monotonic BD

1. The usual axioms and rules for first-order logic
2. KD4 axioms for K (hard belief)
3. S5 axioms for A
4. The Barcan and converse Barcan formulas and necessitation for all modal operators (including F_π and P_π).

5. $\varphi > \varphi$
6. $\vdash ((\varphi_1 \& ... \& \varphi_n) \rightarrow \varphi) \Rightarrow \vdash (((\psi > \varphi_1) \& ... \& (\psi > \varphi_n)) \rightarrow (\psi > \varphi)$
7. $\vdash \varphi \leftrightarrow \psi \Rightarrow \vdash (\varphi > \zeta) \leftrightarrow (\psi > \zeta)$
8. $(\varphi > \zeta) \& (\psi > \zeta) \rightarrow ((\varphi \vee \psi) > \zeta)$
9a. $F\pi(\varphi \rightarrow \psi) \rightarrow (F\pi\varphi \rightarrow F\pi\psi)$ 9b. $P\pi(\varphi \rightarrow \psi) \rightarrow (P\pi\varphi \rightarrow P\pi\psi)$
10a. $\mathcal{D}(\pi) \rightarrow F\pi\mathcal{R}(\pi)$ 10b. $\mathcal{R}(\pi) \rightarrow P\pi\mathcal{D}(\pi)$
11a. $\mathcal{D}(\pi) \rightarrow (F\pi\varphi \leftrightarrow \neg F\pi\neg\varphi)$ 11b. $\mathcal{R}(\pi) \rightarrow (P\pi\varphi \leftrightarrow \neg P\pi\neg\varphi)$
12. $(\mathcal{R}(\pi) \& \varphi) \rightarrow P\pi F\pi\varphi$
Interpreting Complex Programs:
13a. $\mathcal{D}(\pi|\pi') \leftrightarrow (\mathcal{D}(\pi) \vee \mathcal{D}(\pi'))$ 13b. $\mathcal{R}(\pi|\pi') \leftrightarrow (\mathcal{R}(\pi) \vee \mathcal{R}(\pi'))$
14a. $\mathcal{D}(?\varphi) \leftrightarrow \varphi$ 14b. $\mathcal{R}(?\varphi) \leftrightarrow \varphi$
15a. $\mathcal{D}(\pi;\pi') \leftrightarrow (\mathcal{D}(\pi) \& F\pi\mathcal{D}(\pi'))$ 15b. $\mathcal{R}(\pi;\pi') \leftrightarrow (\mathcal{R}(\pi') \& P\pi'\mathcal{R}(\pi))$
16a. $(\mathcal{D}(\pi) \& \varphi) \leftrightarrow \mathcal{D}(?\varphi;\pi)$ 16b. $(\mathcal{R}(\pi) \& \varphi) \leftrightarrow \mathcal{R}(\pi;?\varphi)$
Temporal transitivity:
17a. $\mathcal{D}(\pi_1;(\pi_2;\pi_3)) \leftrightarrow \mathcal{D}((\pi_1;\pi_2);\pi_3)$ 17b. $\mathcal{R}(\pi_1;(\pi_2;\pi_3)) \leftrightarrow \mathcal{R}((\pi_1;\pi_2);\pi_3)$
Uniqueness of next/last state:
18a. $N(\varphi \rightarrow \psi) \rightarrow (N\varphi \rightarrow N\psi)$ 18b. $L(\varphi \rightarrow \psi) \rightarrow (L\varphi \rightarrow L\psi)$
Necessity of past:
19. $\varphi \rightarrow A\varphi$, where φ contains no occurrence of '$\mathcal{D}$', 'N' or 'F'
Connections between temporal and dynamic modalities
20a. $F\alpha\varphi \rightarrow N\varphi$, where α is basic action 20b $P\alpha\varphi \rightarrow L\varphi$, where α is basic action

Appendix B. Proof of the Rational Monotonicity of CE for non-nested language.

Fact 5: Let α, β, γ be modality-free formulae and let T be a set of simple unnested > formulae or strict conditionals.

(ia) $T, \alpha \vdash \beta \Rightarrow T, \alpha \mid\!\approx_1 \beta$	SUPRACLASSICALITY
(ib) $T, \alpha \mid\!\approx_1 \beta \;\&\; \vdash \alpha \leftrightarrow \gamma \Rightarrow T, \gamma \mid\!\approx_1 \beta$	LE
(ii) $T, \alpha \mid\!\approx_1 \beta \;\&\; T, \alpha, \beta \mid\!\approx_1 \gamma \Rightarrow T, \alpha \mid\!\approx_1 \gamma$	CUT
(iii) $T, \alpha \mid\!\approx_1 \beta \;\&\; T, \alpha \mid\!\approx_1 \gamma \Rightarrow T, \alpha, \beta \mid\!\approx_1 \gamma$	CM
(iv) $T, \alpha \mid\!\approx_1 \beta \;\&\; T, \gamma \mid\!\approx_1 \beta \Rightarrow T, \alpha \vee \gamma \mid\!\approx_1 \beta$	OR
(v) $T, \alpha \mid\!\approx_1 \beta$ & not$(T, \alpha \mid\!\approx_1 \neg\gamma) \Rightarrow T, \alpha \& \gamma \mid\!\approx_1 \beta$	RM

Properties (ia) and (ib) stem directly from the definition of $\mid\!\approx_1$.

To show (ii) CUT (and simultaneously (iii) CM) we show that the $T \cup \{\alpha\}$ fixpoints are also $T \cup \{\alpha, \beta\}$ fixpoints by induction along the normalization process. More specifically, we prove by induction that for a given ordering ζ of ANT(T), where $\mathcal{B}$ is the $T \cup \{\alpha\}$ fixpoint on the ordering ζ and $X\delta(\zeta)$ δ-th normalization of X on ordering ζ, $\mathcal{B}\delta(\zeta) \subseteq \|T \cup \{\alpha, \beta\}\|\delta(\zeta) \subseteq \|T \cup \{\alpha\}\|\delta(\zeta)$. for all δ. Observe that for $*(\mathcal{B}, p) \cap \mathcal{B} \neq \emptyset$, then NORM$(\mathcal{B}, p) \subseteq$ NORM$(\|T \cup \{\alpha, \beta\}\|, p) \subseteq$ NORM$(\|T \cup \{\alpha\}\|, p)$. So assume that p in the enumeration ζ is the first proposition such that $*(\mathcal{B}, p) \cap \mathcal{B} = \emptyset$ and that $X\delta{+}1(\zeta) =$ NORM$(X\delta(\zeta), p)$. The inductive hypothesis is that $\mathcal{B}\delta(\zeta) \subseteq \|T \cup \{\alpha, \beta\}\|\delta(\zeta) \subseteq \|T \cup \{\alpha\}\|\delta(\zeta)$. Note that $\mathcal{B}\delta(\zeta) = \mathcal{B}$, since $\mathcal{B}$ is a fixed point.

(a) Suppose now that $*(\|T \cup \{\alpha, \beta\}\|\delta(\zeta), p) \cap \|T \cup \{\alpha, \beta\}\|\delta(\zeta) = \emptyset$. Then if $*(\|T \cup \{\alpha\}\|\delta(\zeta), p) \cap \|T \cup \{\alpha\}\|\delta(\zeta) = \emptyset$, we have $\mathcal{B}\delta{+}1(\zeta) \subseteq \|T \cup \{\alpha, \beta\}\|\delta{+}1(\zeta) \subseteq \|T \cup \{\alpha\}\|\delta{+}1(\zeta)$. So suppose not. Then $*(\|T \cup \{\alpha\}\|\delta(\zeta), p) \cap \|T \cup \{\alpha\}\|\delta(\zeta) \neq \emptyset$. So $*(\|T \cup \{\alpha\}\|1(\zeta), p_1) \cap *(\|T \cup \{\alpha\}\|2(\zeta), p_2) \cap \ldots \cap *(\|T \cup \{\alpha\}\|\delta{-}1(\zeta), p\delta) \cap *(\|T \cup \{\alpha\}\|\delta(\zeta), p) \neq \emptyset$, for p_j such that $\zeta(p_j) = j$. Because T entails no nested conditionals and α is > free, $(T \cup \{\alpha\})(n, \zeta) \vdash \varphi > \psi$ implies $T \cup \{\alpha\} \vdash \varphi > \psi$. So then where $T \vdash p_j > \varphi_{jn}$ and where $T \vdash p > \psi_m$, for each $j \leq \delta$ and each n and m, $(\bigcap j \leq \delta(\bigcup\{\varphi_{jn}\})) \cap (\bigcup\{\|\psi_m\|\}) \cap \|T \cup \{\alpha\}\| \neq \emptyset$. But $*(\|T \cup \{\alpha, \beta\}\|1(\zeta), p_1) \cap *(\|T \cup \{\alpha, \beta\}\|2(\zeta), p_2) \cap \ldots \cap *(\|T \cup \{\alpha, \beta\}\|\delta{-}1(\zeta), p\delta) \cap *(\|T \cup \{\alpha, \beta\}\|\delta(\zeta), p) = \emptyset$.

Again since T entails no nested conditionals and α and β are > free, $(T \cup \{\alpha, \beta\})(n, \zeta) \vdash \varphi > \psi$ implies $T, \alpha, \beta \vdash \varphi > \psi$. So since p is the first proposition where $*(\mathcal{B}, p) \cap \mathcal{B} = \emptyset$, $(\bigcap_{j \leq \delta}(\|\bigcup\{\varphi_{jn}\}\|) \cap (\bigcup\{\|\psi_m\|\}) \cap \|T \cup \{\alpha, \beta\}\| = \emptyset$. So then $(\bigcap_{j \leq \delta}(\bigcup\{\varphi_{jn}\})) \cap (\|\bigcup\{\psi_m\}\|) \cap \|T \cup \{\alpha\}\| \models \neg\beta$. Thus, $\|T \cup \{\alpha\}\|_{\delta+1(\zeta)} \models \neg\beta$, which contradicts our hypothesis that every $T \cup \{\alpha\}$ fixpoint verifies β.

(b) So now assume $*(\|T \cup \{\alpha, \beta\}\|_{\delta(\zeta)}, p) \cap \|T \cup \{\alpha, \beta\}\|_{\delta(\zeta)} \neq \emptyset$. Then $*(\|T \cup \{\alpha\}\|_{\delta(\zeta)}, p) \cap \|T \cup \{\alpha\}\|_{\delta(\zeta)} \neq \emptyset$. Since $NORM_1$ is a monotonic decreasing function in this case, we have $\|T \cup \{\alpha, \beta\}\|_{\delta+1(\zeta)} \subseteq \|T \cup \{\alpha\}\|_{\delta+1(\zeta)}$. It now remains to show also $\mathcal{B}_{\delta+1(\zeta)} \subseteq \|T \cup \{\alpha, \beta\}\|_{\delta+1(\zeta)}$. Suppose not; then $\exists w'$ $w' \in \mathcal{B}$, but $w' \notin *(*(\|T \cup \{\alpha, \beta\}\|_{\delta(\zeta)}, p)$ and $w' \models p$. But since $\mathcal{B} \subseteq \|T \cup \{\alpha\}\|_{\delta+1(\zeta)}$, $w' \in NORM_1(\|T \cup \{\alpha\}\|_{\delta(\zeta)}, p)$. So $\exists w'' \in \|T \cup \{\alpha\}\|_{\delta(\zeta)}$ such that $w' \in *(w'', p)$. But given $w'' \models T$ and β is > free, $*(w'', p) \subseteq *(\|T \cup \{\alpha, \beta\}\|_{\delta(\zeta)}, p)$, which contradicts our assumption that $w' \notin *(*(\|T \cup \{\alpha, \beta\}\|_{\delta(\zeta)}, p)$. So we have shown $\mathcal{B}_{\delta+1(\zeta)} \subseteq \|T \cup \{\alpha, \beta\}\|_{\delta+1(\zeta)} \subseteq \|T \cup \{\alpha\}\|_{\delta+1(\zeta)}$. The limit case is routine. Thus, we have shown that, where λ is the stage at which the revision process for all three sets reaches a fixed point, $\mathcal{B} = \mathcal{B}_{\lambda(\zeta)} \subseteq \|T \cup \{\alpha, \beta\}\|_{\lambda(\zeta)} \subseteq \|T \cup \{\alpha\}\|_{\lambda(\zeta)} = \mathcal{B}$-- which is what we desired to prove.

To show (iv), assume that $T, \alpha \mid\!\approx \beta$ and $T, \gamma \mid\!\approx \beta$. Show $T, \alpha \vee \gamma \mid\!\approx \beta$. Suppose w is a world that survives the normalization procedure on $T \cup \{\alpha \vee \gamma\}$; i.e. $w \in \mathcal{B}$, where $\mathcal{B}$ is any $T \cup \{\alpha \vee \gamma\}$ fixed point. Then $w \models \alpha$ or $w \models \gamma$. Suppose $w \models \alpha$. If w is an element of every $T \cup \{\alpha\}$ fixpoint, then $w \models \beta$ and we can then show that $\mathcal{B} \models \beta$ (see Asher and Morreau 1991). So assume there is a $\mathcal{B}'$ that is a $T \cup \{\alpha\}$ fixpoint but $w \notin \mathcal{B}'$. Then for some p_0 and some enumeration ζ. $w \notin NORM(\|T \cup \{\alpha\}\|_{\delta(\zeta)}, p_0)$ and $w \in \|T \cup \{\alpha\}\|_{\delta(\zeta)}$. By hypothesis, $*(\|T \cup \{\alpha\}\|_{\delta(\zeta)}, p_0) \cap \|T \cup \{\alpha\}\|_{\delta(\zeta)} \neq \emptyset$. Further $w \models p_0$ but $\neg\exists w' \in \|T \cup \{\alpha\}\|_{\delta(\zeta)}$ such that $w \in *(w', p_0)$. So I claim that for some ψ, $w \models p_0$ & ψ and $\|T \cup \{\alpha\}\|_{\delta(\zeta)} \models p_0 > \neg\psi$. Suppose not, and that $\|T \cup \{\alpha\}\|_{\delta(\zeta)} \models p_0 > \neg\psi_j$ for all ψ_j such that $w \models p_0 > \psi_j$. Then $p_0 \cap (\|\bigcup\{\psi_j\}\|) \cap \|T \cup \{\alpha\}\|_{\delta(\zeta)} \neq \emptyset$. But then we can show that there is a $w' \in W_{CAN}$ such that $w' \in \|T \cup \{\alpha\}\|_{\delta(\zeta)}$ and $w \in *(w', p_0)$. So it must be the case that: or some ψ, $w \models p_0$ & ψ and $\|T \cup \{\alpha\}\|_{\delta(\zeta)} \models p_0 > \neg\psi$. Since w is a $T \cup \{\alpha \vee \beta\}$ survivor world, however, then $\|T \cup \{\alpha \vee \beta\}\|_{\delta(\zeta)} \models p_0 > \neg\psi$. Because we have no nested conditionals in T or α, $\|T \cup \{\alpha\}\|_{\delta(\zeta)} \models p_0 > \neg\psi$ implies $T \cup \{\alpha\} \models p_0 > \neg\psi$. And because α is modal free, $\alpha \cup \{\neg(p_0 > \neg\psi)\}$ is consistent, and so $T \cup \{\alpha\} \models p_0 > \neg\psi$ implies $T \models p_0 > \neg\psi$ and hence $T \cup \{\alpha \vee \gamma\} \models p_0 > \neg\psi$, which contradicts our hypothesis that w is a $T \cup \{\alpha \vee \gamma\}$ survivor world. Thus, if w is not a $T \cup \{\alpha\}$ survivor world then $w \models \neg\alpha$. So assume $w \models \gamma$. By a similar argument we show that w must then be a $T \cup \{\gamma\}$ survivor world. So whether $w \models \alpha$ or $w \models \gamma$, $w \models \beta$.

Finally, to show (v). Assume $T, \alpha \mid\!\approx \beta$ and not($T, \alpha \mid\!\approx \neg\gamma$). Then there is at least one $T \cup \{\alpha\}$ fixpoint $\mathcal{B}$ and there is a $w \in \mathcal{B}$ such that $w \models \alpha$ & β & γ. Now consider any $T \cup \{\alpha \,\&\, \gamma\}$ fixpoint $\mathcal{B}'$. I claim $w \in \mathcal{B}'$ and so $\mathcal{B}' \models \beta$. Suppose not. Then there is a $\mathcal{B}''$ such that $w \notin \mathcal{B}''$ and so for some p_0 and some ordering ζ, $w \notin NORM(\|T \cup \{\alpha \,\&\, \gamma\}\|_{\delta(\zeta)}, p_0)$ and so $*(\|T \cup \{\alpha \,\&\, \gamma\}\|_{\delta(\zeta)}, p_0) \cap \|T \cup \{\alpha \,\&\, \gamma\}\|_{\delta(\zeta)} \neq \emptyset$. $\neg\exists w' \in \|T \cup \{\alpha \,\&\, \gamma\}\|_{\delta(\zeta)}$ such that $w \in *(w', p_0)$. So again we claim $w \models p_0$ & ψ and $\|T \cup \{\alpha \,\&\, \gamma\}\|_{\delta(\zeta)} \models p_0 > \neg\psi$ for reasons similar to those given in the proof OR. But again this implies, given that T is non-nested and α and γ are modal free that $T, \alpha \models p_0 > \neg\psi$, and this contradicts our assumption that w is an element of a $T \cup \{\alpha\}$ fixpoint. Since we have verified (v) and (v) implies (iii), we have an alternative proof of (iii) as well.□

Thus, the consequence relation $\mid\!\approx$ relative to a fixed set of conditionals defined above is a rational consequence relation.[4] For this restricted language, CE has the same nice metatheoretic properties discovered by Lehmann as 1-entailment (Pearl 1990), the Rational Closure of Lehmann (1989), the S4.3 models of Boutilier and Lamarre, and the $\vdash_\pi$ relation of Benferhat, Dubois and Prade (1992).

[4] We cannot show, however, in general that every ranked model used to define the rational consequence relation can be generated with a theory in CE.

Appendix C. Representation Theorem for Commonsense Entailment. (Proof of Fact 6)

The basic notion of the proof theoretic representation of $\models_1$ is an *extension* and an *extension sequence*. I define an *extension sequence* of a finite set Γ *relative to* $\vdash_{BD}$ (the monotonic core entailment notion) and a function ζ from ANT(Γ) $\rightarrow$ N as follows:

Definition: Let Γ be a finite set of sentences, and let m be the number of sentences in ANT(Γ).

$Ext^0(\Gamma, \zeta) = \Gamma$

$Ext^{\beta+1}(\Gamma, \zeta) = Ext^{\beta}(\Gamma, \zeta) \cup \{\psi \rightarrow \varphi: Ext^{\beta}(\Gamma, \zeta) \vdash \psi > \varphi\}$, where $\zeta(\psi) = n + 1$ and $\beta+1 = km + n + 1$, for some natural number k, provided the following is consistent: $Ext^{\beta}(\Gamma, \zeta) \cup \{\varphi: Ext^{\beta}(\Gamma, \zeta) \vdash \psi > \varphi\}$.

$= Ext^{\beta}(\Gamma, \zeta)$, otherwise.

$Ext^{\lambda}(\Gamma, \zeta) = \cup_{\beta < \lambda} Ext^{\beta}(\Gamma, \zeta)$ for limit ordinals λ.

Lemma C1: Every extension sequence of a set of premises in L_{BD} has a fixpoint.
Proof: By induction on the ordinals.

Whenever Ext reaches a fixed point for a given Γ and given ζ, then I say we have a Γ extension.

Lemma C2: For finite Γ, $|ANT(\Gamma)| = n$ with maximal depth of conditionals m, for every ordering ζ of ANT(Γ), an extension of Γ is obtained by nm + 1.

Proof sketch: Suppose $|ANT(\Gamma)| = n$ with maximal depth of conditionals m. But suppose for some ordering ζ of ANT(Γ), $Ext^{nm+1}(\Gamma, \zeta) \neq Ext^{nm}(\Gamma, \zeta)$. Then for some $\psi \in ANT(\Gamma)$, it's not the case that $\{\psi \rightarrow \varphi: Ext^{nm}(\Gamma, \zeta) \vdash \psi > \varphi\} \subseteq Ext^{nm}(\Gamma, \zeta)$. Given that $|ANT(\Gamma)| = n$, we may assume $\zeta(\psi) = k \leq n$. By hypothesis there is a formula $\psi \rightarrow \varphi \in Ext^{nm+1}(\Gamma, \zeta)$, $\psi \rightarrow \varphi \notin Ext^{km}(\Gamma, \zeta)$, and so $Ext^{nm}(\Gamma, \zeta) \vdash \psi > \varphi$ and $Ext^{km-1}(\Gamma, \zeta) \nvdash \psi > \varphi$. In view of the construction, we have minimally $Ext^{km-1}(\Gamma, \zeta) \vdash (\psi > (\psi > \varphi))$ & ψ, so that on the next turn of examining ψ, $Ext^{km-1}(\Gamma, \zeta) \vdash \psi > \varphi$. By reasoning in this way for each kpth stage for each p, $1 \leq p \leq m$, we see that there are at least $\delta_1, \ldots \delta_m$ such that $\Gamma \vdash \delta_1 > (\delta_2 > \ldots > (\delta_m > \varphi)\ldots)$. This now contradicts our assumption that the maximal nesting depth of Γ is m.□

We now have the main lemma linking the semantic notion of normalization with extension sequences.

Lemma C3 Let Γ be a finite set of sentences of L_{BD}. For all orderings ζ of ANT(Γ) and all natural numbers β,
$Ext^{\beta}(\Gamma, \zeta) \vdash \varphi$ iff $\Gamma_{\beta}(\zeta) \models \varphi$.

Proof Sketch: By induction on the natural numbers.
Base Case: obvious from the completeness theorem for CE, given that $\Gamma(0, \zeta) = \Gamma$ and $\sigma_0(\zeta) = W + \Gamma$.
Inductive Step: Assume an ordering ζ of ANT(Γ) and assume that $\Gamma(n, \zeta) \vdash \varphi$ iff $\Gamma_{\beta}(\zeta)$. Show that $Ext^{\beta+1}(\Gamma, \zeta) \vdash \varphi$ iff $\Gamma_{\beta+1}(\zeta) \models \varphi$.

$\Rightarrow$ Suppose $Ext^{\beta+1}(\Gamma, \zeta) \vdash \varphi$. I show $\Gamma_{\beta+1}(\zeta) \models \varphi$. Suppose $\Gamma_{\beta+1}(\zeta) \not\models \varphi$. Then there is a world w such that: $w \in \Gamma_{\beta+1}(\zeta)$ and $w \models \neg\varphi$, where $\Gamma_{\beta+1}(\zeta) = NORM(\Gamma_{\beta}(\zeta), \|\psi\|)$, for $\zeta(\psi) = n+1$, where $\beta + 1 = km + n + 1$, for m the size of ANT(Γ), and k some natural number. By the definition of NORM, $\Gamma_{\beta+1}(\zeta) \subseteq \Gamma_{\beta}(\zeta)$. So $w \in \Gamma_{\beta}(\zeta)$. By the inductive hypothesis, $\Gamma_{\beta}(\zeta) \models \varphi$ iff $Ext^{\beta}(\Gamma, \zeta) \vdash \varphi$. So if $Ext^{\beta}(\Gamma, \zeta) \vdash \psi > \delta$, then $\Gamma_{\beta}(\zeta) \models \psi > \delta$ and so $w \models \psi > \delta$. Since by hypothesis $w \in NORM(\Gamma_{\beta}(\zeta), \|\psi\|)$, $w \not\models \psi$ or $w \models \psi$ & δ' for each δ' such that $\Gamma_{\beta}(\zeta) \models \psi > \delta$. So $w \models \psi \rightarrow \delta'$ for each

δ' such that $\Gamma_\beta(\zeta) \models \psi > \delta'$. By the inductive hypothesis, $w \models \psi > \delta'$ for each δ' such that $\mathrm{Ext}^\beta(\Gamma, \zeta) \vdash \psi > \delta'$. So then $w \models \Gamma(n+1,\zeta)$, since $\mathrm{Ext}^{\beta+1}(\Gamma, \zeta) = \mathrm{Ext}^\beta(\Gamma, \zeta) \cup \{p \rightarrow \delta: \Gamma(n, \zeta) \vdash p > \delta\}$, where $\mathrm{Ext}^\beta(\Gamma, \zeta) \cup \{\delta: \Gamma(n, \zeta) \vdash p > \delta\}$ is consistent. By the soundness of CE deduction, then $w \models \varphi$, which contradicts our hypothesis.

$\Leftarrow$ Suppose $\Gamma_{\beta+1}(\zeta) \models \varphi$. We now show $\mathrm{Ext}^{\beta+1}(\Gamma, \zeta) \vdash \varphi$. Suppose $\mathrm{Ext}^{\beta+1}(\Gamma, \zeta) \nvdash \varphi$. By the inductive hypothesis $\mathrm{Ext}^\beta(\Gamma, \zeta) \vdash \beta$ iff $\Gamma_\beta(\zeta) \models \beta$. So our hypothesis and the definition of $\mathrm{Ext}^{\beta+1}(\Gamma, \zeta)$ entail that $\Gamma_\beta(\zeta) \subset \Gamma_{\beta+1}(\zeta)$ and that $\Gamma_\beta(\zeta) \not\models \varphi$. Thus $*(\Gamma_\beta(\zeta), p) \cap \Gamma_\beta(\zeta) \neq \emptyset$, and by the inductive hypothesis $\mathrm{Ext}^\beta(\Gamma, \zeta)$ is consistent with $\{\varphi: \Gamma(n, \zeta) \vdash \psi > \varphi\}$. In the canonical model the worlds of which $\Gamma_{\beta+1}(\zeta)$ is a subset, $\Gamma_{\beta+1}(\zeta) = \{w \in O: \Gamma_\beta(\zeta) \models p > \psi \rightarrow (w \not\models p \vee w \models p \,\&\, \psi)\} = \{w \in O: w \models \Gamma(n, \zeta) \,\&\, \Gamma(n, \zeta) \vdash p > \psi \rightarrow (w \not\models p \vee w \models p \,\&\, \psi)\} = \{w \in O: \mathrm{Ext}^\beta(\Gamma, \zeta) \subseteq w \,\&\, \mathrm{Ext}^\beta(\Gamma, \zeta) \vdash p > \psi \rightarrow (\neg p \in w \vee p \,\&\, \psi \in w)\} = \{w \in O: \mathrm{Ext}^\beta(\Gamma, \zeta) \subseteq w \,\&\, \mathrm{Ext}^\beta(\Gamma, \zeta) \vdash p > \psi \rightarrow (p \rightarrow \psi \in w)\} = \{w \in O: \mathrm{Ext}^{\beta+1}(\Gamma, \zeta), \zeta) \subseteq w)\}$. Since O contains all the maximal $\vdash_{CE}$-consistent sets and φ is an element of every w in the set above, $\mathrm{Ext}^{\beta+1}(\Gamma, \zeta) \vdash \varphi$, contrary to hypothesis.

Corollary C4: For all Γ extensions $\mathcal{A}$, there is a Γ fixpoint $\mathcal{B}$ such that $\mathcal{A} \vdash \varphi$ iff $\mathcal{B} \models \varphi$ and for all Γ fixpoints $\mathcal{B}$ there is a Γ extension $\mathcal{A}$ such that $\mathcal{A} \vdash \varphi$ iff $\mathcal{B} \models \varphi$.

Definition: $\Gamma \mid\sim \varphi$ iff for all Γ extensions $\mathcal{A}$, $\mathcal{A} \vdash \varphi$.

Corollary C5: $\Gamma \mid\sim \varphi$ iff $\Gamma \models\approx \varphi$.

The notion of an extension exploits consistency checks and $\vdash$. Thus, for propositional commonsense entailment we have a decidable proof procedure for finite theories. The following theorem gives us a necessary and sufficient method to determine as to whether there is a fixpoint defined from a finite theory that verifies A (or $\neg$A). It is the essential element to encoding within the monotonic conditional logic of BD a CE notion of nonmonotonic consequence. For the rest of the paper we shall simply work with finite theories, and so we will consider only finitary extensions.

Definition. $\beth(\Gamma,\zeta) = \rho$ iff there is a subset $\{\Psi_1,...,\Psi_n\}$ of ANT(Γ), and L_{BD} formulas $\Phi_1,...,\Phi_n$ such that

(i) ρ is of the form $(\Psi_1 \rightarrow \Phi_1) \,\&... \&\, (\Psi_n \rightarrow \Phi_n)$

(ii) for each $j < n$, $\zeta(\Psi_j) < \zeta(\Psi_{j+1})$.

(iii) For each $j < n$, $\Gamma \cup \Phi_1$, $\Gamma \cup \{\Psi_1 \rightarrow \Phi_1\} \cup \{\Phi_2\}, ... \Gamma \cup \{\Psi_1 \rightarrow \Phi_1\} \cup \ldots \cup \{\Psi_j \rightarrow \Phi_j\} \cup \{\Phi_{j+1}\}$ are all consistent.

(iv) Φ_{j+1} is the first formula θ (in some fixed enumeration of the language) such that for all φ such that $\Gamma \cup \{\Psi_1 \rightarrow \Phi_1, \ldots, \Psi_j \rightarrow \Phi_j\} \vdash \Psi_{j+1} > \varphi$, $\vdash \theta \rightarrow \varphi$.

(v) There is no formula $\delta \in \mathrm{ANT}(\Gamma)\backslash\{\Psi_1, \ldots, \Psi_n\}$ and no Ψ_j such that $\zeta(\delta) < \zeta(\Psi_{j+1})$ and $\Gamma \cup \{(\Psi_1 \rightarrow \Phi_1), ... (\Psi_j \rightarrow \Phi_j)\} \cup \{\varphi: \Gamma, \{\Psi_1 \rightarrow \Phi_1, \ldots, \Psi_j \rightarrow \Phi_j\} \vdash \delta > \varphi\}$ is consistent.

(vi) There is no formula $\delta \in \mathrm{ANT}(\Gamma)\backslash\{\Psi_1, \ldots, \Psi_n\}$ and $\Gamma \cup \{\beth(\Gamma,\zeta)\} \cup \{\varphi: \Gamma, \beth(\Gamma,\zeta) \vdash \delta > \varphi\}$ is consistent.

Theorem C6: Let Γ be a finite theory. There is a Γ extension $\mathcal{A}$ such that $\mathcal{A} \vdash \psi$ iff there is an ordering ζ such that $\Gamma, \beth(\Gamma,\zeta) \vdash \psi$.

Proof Sketch: $\Rightarrow$ Suppose that there is a Γ extension $\mathcal{A}$ such that $\mathcal{A} \vdash \psi$. $\mathcal{A}$ is the fixpoint of an extension sequence defined relative to some ordering ζ. Since Γ is finite, by lemma C2 the extension sequence is finite and consists of at most km + 1 steps for $|\mathrm{ANT}(\Gamma)| = k$ and depth of nesting m. Let us suppose that the formulas in ANT(Γ) defining the extension sequence in order are $\Psi_1, \ldots, \Psi_k$. Assume $\Gamma(1, \zeta) \neq \Gamma$. By definition, $\Gamma(1, \zeta) = \Gamma \cup \{\Psi_1 \rightarrow \Phi_1: \Gamma \vdash \Psi_1 > \Phi_1 \,\&\, \vdash\Phi_1 \rightarrow \varphi$, for all φ such that $\Gamma \vdash \Psi_1 > \varphi\}$, provided Γ is consistent with Φ_1. Let us assume Γ is consistent with Φ_1;

otherwise, $\Gamma(1, \zeta) = \Gamma$. For every $\gamma \in \Gamma(1, \zeta)$, $\Gamma \cup \{\Psi_1 > \Phi_1, \Phi_1 \vee \neg\Psi_1\} \vdash \gamma$. Using the deduction theorem for $\vdash_{BD}$, for every $\gamma \in \Gamma(1, \zeta)$, $\Gamma \vdash (\Psi_1 > \Phi_1)$ & $((\Psi_1 \rightarrow \Phi_1) \rightarrow \gamma)$.

Now consider Ψ_2. If $\Gamma(1, \zeta) \vdash \Psi_2 > \Phi_2$, then $\Gamma \vdash (\Psi_1 > \Phi_1)$ & $((\Psi_1 \rightarrow \Phi_1) \rightarrow (\Psi_2 > \Phi_2))$. Assume now that Φ_2 is consistent with $\Gamma \cup \{\Psi_1 \rightarrow \Phi_1\}$. Otherwise, $\Gamma(2, \zeta) = \Gamma(1, \zeta)$. Continuing on in this way to the end of the extension sequence to the fixpoint $\mathcal{A}$, we get the desired formula as well as the satisfaction of conditions (i) -(vi) of the definition of $\beth(\Gamma,\zeta)$. Since the theory is finite the sequence is finite, and so the formula is a formula of L_{BD}.

$\Leftarrow$ Conversely, suppose that Γ, $(\Psi_1 \rightarrow \Phi_1)$ &... & $(\Psi_n \rightarrow \Phi_n) \vdash \psi$, for some ζ and some $\Psi_1 \ldots$, $\Psi_n \in$ ANT(Γ) with conditions (i) -(vi) met. Then by the definition of a Γ extension, there is an extension $\mathcal{A} = \Gamma(m, \zeta)$ such that $\mathcal{A} \vdash \psi$.

Discussion: The first clause gives us an encoding of a normalization chain under a given ordering. First, we exploit the > conditional involving the first antecedent and from the normalized result we get $\Psi_1 \rightarrow \Phi_1$. From this we deduce a conditional involving the second antecedent in the ordering, which we exploit in a second normalization. The combination of the first and second normalization yields a state in which $((\Psi_1 \rightarrow \Phi_1)$ & $(\Psi_2 \rightarrow \Phi_2))$; we then use this to look at the conditionals involving Ψ_3 and so on. The other constraints of the definition of $\beth$ ensure that normalization with $\Psi_1, \ldots, \Psi_n$ is successful and exhaustive. Thus, to show $\Gamma \models\!\approx \psi$ for a Γ with ANT(Γ) finite, it suffices to give a proof that there is a formula of the sort defined by $\beth$ with ψ as a conclusion for every ordering of ANT(Γ).

Corollary C7: If Γ is a finite theory then $\Gamma \mid\!\sim \psi$ iff for all orderings ζ of ANT(Γ), $\Gamma, \beth(\Gamma,\zeta) \vdash \psi$.

We do not need anything stronger than embedded conditionals and a consistency check on the formulas they contain to calculate nonmonotonic consequence in CE. For the simple inference patterns given in the literature, these conditionals are simple and easily verified in $\vdash_{BD}$. $\models\!\approx_1$ is equivalent to a proof of a conjunction of nested conditionals of a certain form in the basic monotonic logic together with a number of consistency checks. This elucidates the $\models\!\approx_1$ relation and this also means that automatic proofs of nomonotonic inferences in at least the simplest cases are possible.